

Bin Picking and Uncertainty Modelling of Generalized Cylindrical Object using Sensor Fusion

*

Shraddha Arora^{a,*,*}

, Anshu Khurana^b Shams Tabrez Siddiqui^c,
Chaman Kumar^d, S. Dutta Roy^e

^a Assistant professor (Sel. Grade), Computer Science
and Engineering department, The NorthCap
University, Gurugram-122017

^b Assistant Professor Department of Artificial
Intelligence and Data Science, Maharaja Agrasen
Institute of Technology, Delhi-110086

^c Department of Computer Science, College of
Computer Science and Information Technology, Jazan
University, Jazan, 45142, Saudi Arabia

^d IIMT college of Engineering Greater Noida,
Department of Information Technology

^e Department of Electrical Engineering, Indian
Institute of Technology Delhi

posed solution introduces efficient O-RANSAC-based algorithms, specifically designed to leverage range data obtained from a laser scanner [1], hence eliminating unseen issues. This utilization of range data is instrumental in facilitating the picking and placing of objects within the bin. The research in this paper rigorously tests the effectiveness of the proposed algorithm across various scenarios, encompassing real-time operations, obscured placements, generalized situations, and instances involving featureless objects. The model was tested for multiple and achieved an Accuracy of 94% in the pose estimation for the pickup of large cylindrical pellets and 74% for general cylindrical objects. This Accuracy supersedes the generalized objects when compared to state-of-the-art methods. This paper has also demonstrated reduced errors in the estimation of parameters when compared to the deterministic approach. The quantitative and qualitative analysis shows that the proposed model could be used to reduce manual work.

Keywords: Bin Picking, Pose estimation, autonomous robotics, Sensor Fusion, Computer Vision

Abstract. Automation is a necessity in the changing digital era. It helps to increase the efficiency of any process requiring constant manual work. The automation process faces difficulties whenever an object arrangement task involves both texture and feature loss. The more challenging phase arises when these objects were previously unseen and have unknown dimensions, rendering conventional methods inapplicable. To overcome the existing challenges, we proposed an approach for random bin picking using sensor fusion of image and range sensor data. This innovative approach helps improve performance by handling difficulties encountered when handling objects that lack distinctive features, texture, or unique characteristics, especially when positioned in an obstructed manner within a storage bin. The pro-

1. INTRODUCTION

The fourth Industrial Revolution, also known as Industry 4.0, is directly tied to the rise of industrial automation. Industrial automation is carried out mainly to reduce human intervention and increase the efficiency of the industrial process. Random bin picking using robots is considered the holy grail of robotic vision. The robot has to locate a part in free space, in an unstructured environment where the parts keep shifting positions and orientations every time a part is removed from the bin. Complications arise when many distinct, feature-less and texture-less objects (such as

*

**chaudhary.shraddha18@gmail.com

*

**chaudhary.shraddha18@gmail.com

black cylindrical items) are randomly piled up in the bin, with arbitrary orientations and heavy occlusion. To achieve this task, a set of sensors with complementary properties (a camera and a range sensor) are used for pose estimation in heavy occlusion, accordingly orienting the manipulator gripper to pick up items.

1.1. Our Contribution

There are two significant contributions of this paper. The first is the introduction of a probabilistic approach in determining the position and orientation of feature-less and texture-less cylindrical pellets in a cluttered bin. An optimized approach of RANSAC paradigm is introduced for finding the best-fitting cylinder on the point cloud of pellet. The results are compared with the existing deterministic approach of [2]. The results are further improved by implementing MSAC for parameter estimation in the already proposed approach. Uncertainty in the estimated pose is analyzed and used to show the robustness of the system. Next, we further extend the work to general cylindrical objects of different dimensions and show the robustness of the proposed approach in bin-picking applications.

2. Related Work

In the past few decades, a good amount of work has been carried in the vision-based bin picking system. A category of industrial bin picking system as in [3],[2] are techniques which evaluate features from the raw sensory data and evaluates using a classification algorithm trained on templates of primitive shapes to estimate parameters. Since the system in our case needs to be general, and for unseen objects, template comparison is ruled out. The author [4] describes a multi-view robotic grasping system [5] in this paper. A multi-step approach involving RANSAC, ICP, and genetic algorithm-based registration was used for 6-DOF posture estimation. Disruption detection enhanced grasping efficiency in real-world trials. The author [6] proposed a deep learning-based method for point cloud-based posture estimation, combining instance segmentation and pose prediction [7]. A physically simulated engine generated synthetic datasets, and extensive tests validated the network's efficacy in crowded and obstructed scenarios.

In this [8] the author discusses the uncertainty propagation connection in the camera's image space, the laser scanner's 2D space, the Cartesian space in world

coordinates, and the tool space. Vision sensor was a critical sensor for localization, hence, performance maps were required. The author in [9] employed a deep convolutional neural network model trained on 16,430 annotated depth pictures created synthetically in a physics simulator to predict grasp points directly without object segmentation [10]. The author describes a robot control architecture [11] comprising of touch and visual sensors integrated into a commercially available bin picking device. Tactile sensing augments vision-based picking systems, allowing for in-hand item recognition, which is especially effective for difficult-to-recognize objects such as those with translucent or bright surfaces. Instead of immediately lifting wire harnesses, the author [12] proposes grasping and extracting the object in a circle-like trajectory until it is untangled. The approach yields an overall success rate of 89.36%, compared to 42.23% in the baseline. The author [13] offered a flexible robotic system for kitting and assembly jobs that does not require jigs or commercial tool changers in this paper [14]. In study [15] a unique reactive planning method based on the contact information to make the unwrapping process effective and resilient on pallets with unpredictable location or shape. The authors [16] proposed an integrated approach for object recognition and trajectory prediction using multi-view representations of LiDAR returns and camera images. Their model combines Bird's-Eye View (BEV) features from historical LiDAR data with high-definition maps, augmented by additional LiDAR Range-View (RV) features. The resulting fused features are processed in an end-to-end trainable network for final detections and trajectories, achieving efficient LiDAR-camera fusion. The author [17] reviews 3D object identification algorithms, commonly used sensors, and datasets in autonomous vehicles (AVs). They categorize current efforts into monocular, point cloud-based, and fusion approaches based on sensor modalities. The outcomes of these works are summarized, and research gaps and future objectives are identified. AV perception systems, often incorporating machine learning (e.g., deep learning), convert sensory data into semantic knowledge for autonomous driving. Given a packed heap of fresh items of various types, the author provides a robust solution for stable and collision-free grab planning. The author [18] proposes a collision-free grab planning solution for a packed heap of fresh items. Their approach incorporates a unique grab pose rating algorithm and a pose refining method within the grasp planning pipeline. This ensures stable contact between gripper-fingers and the

target item. Due to the densely packed arrangement of items, traditional grasp planning algorithms may struggle to find suitable grip positions. Remarkably, their approach successfully clears up to 14 items using a physical robot equipped with a two-fingered parallel-jaw gripper and a depth sensor. The author's representation in reference [19] to pick up featureless cylinders using deep learning and assembly with a combination of visual and physical input. YOLO object detection was employed to detect the item for robotic pick up. Even in busy surroundings, the item could be identified. The assembly of a cylindrical item in a hole with a little clearance is also covered. The approach employs hybrid control, which consists of position control along the assembly plane and force control along the normal to the plane direction. The author's [3] [20] single-picture hand-eye calibration approach enables collision-free grab planning for robotic cube picking with a monocular camera. By segmenting images from specific camera postures, the method recovers object information despite low contrast or varying illumination. Edge information guides posture calculation, and a directional thresholding technique estimates an extra edge. The 3D edge data computes the item's attitude for successful robotic pick-up. The paper [52] proposes an improved YOLO framework for detecting small industrial objects. By integrating receptive-field attention and multi-scale feature fusion, the model significantly enhances feature representation and detection accuracy, particularly in complex and cluttered industrial environments.

A general object's shape detection technique from point cloud data can be formulated using RANSAC [21]. These techniques, however being general are not fast because of their iterative nature. A faster way for RANSAC is also developed by [22], but the work basically concentrates on sleeping pellets and its matching with templates, our work is more complete and robust because of the use of sensor fusion. In [8] the major work was concentrated on using RANSAC on multiple primitive shapes like spheres, cylinder, cones, and tori. The objective is very broad, checks support for each and every shape and end up consuming more time. The efficiency is affected in such a system; our system is optimized for cylinders and is fast and accurate. [23] uses the geometric primitives of cylinders in visual odometry and uses PCA to classify planes and cylinders, and using RANSAC to estimate cylinder parameters. It also accounts the uncertainty and accordingly optimizes the pose [24]. Our work is most similar to this but uses sensor fusion for the planar surfaces, MSAC for cylinders and models uncertainty in the robotic pickups. The study [50] presents a method that combines instance segmentation with 3D point cloud analysis to accurately estimate object poses in cluttered industrial bins. The approach improves recognition and

localization of partially occluded or overlapping objects, enhancing the efficiency and precision of robotic bin-picking operations. The paper [51] introduces a pose estimation framework that utilizes sparse convolutional networks to efficiently process 3D point clouds. By focusing on sparse geometric features, the method achieves faster computation and higher accuracy in determining object poses within cluttered bin-picking environments.

In [25], Torr et al. present a modification to the RANSAC paradigm as M-estimator Sample Consensus (MSAC). In MSAC, consensus set is determined using M-estimators instead of the ranking based used in RANSAC. Some authors also focussed on energy efficiency presenting solution like Energy Efficient Dynamic and Adaptive State-Based Scheduling (EDASS) [26] and some authors worked on crowd sensing using IOT [27]. Recent works for general objects are done using Deep Learning (DL) and reinforcement learning (RL); these techniques, however, requires hours of training and hence are not directly deployable for new objects. [28] proposed a self-learning process and utilizes the vector obtained on passing through Convolutional Neural Network (CNN) [29] for comparison. CNN based methods with multi-view for 6D pose estimation [30] [31] has achieved great results, however, is again a data-driven self-supervised approach requiring multiple scans from different positions and is not optimized for 1D pose estimation, which is the advantage of our system. [32] [33] presents CNN based system which assigns objects affordances with different grasping primitives, the amount of computing power required for such a system is very high and relies on multiple GPUs. State of the art for our task with texture-less and feature-less object is [34] which achieves 93% pickup success in 1200 attempts. This work is based on a deterministic estimation of the pellet pose and requires prior knowledge of dimensions of the object to be picked. A better sparse Bayesian learning (SBL) scheme is suggested in the research paper [35] for use in high-precision data modeling for structural Health Monitoring (SHM) applications. By adding enhancements to the conventional SBL technique, the authors address the difficulty of effectively modeling SHM measurements by improving both its generalization ability and predictive Accuracy. The study [36] presents a brand-new method for system reliability analysis that makes use of junction tree algorithms and Bayesian networks (BN). The suggested framework provides a flexible and scalable approach to analyze the reliability of complex systems more accurately and computationally efficiently. [37] is a research that presents a new Bayesian dynamic linear model (DLM) framework for missing data imputing and structural health monitoring (SHM) data forecasting during typhoon events. The authors provide a Bayesian method that uses dynamic linear modeling approaches to capture time-varying patterns and forecast upcoming observations while taking temporal dependencies and uncertainties in SHM data into account.

3. SENSOR FUSION AND CONFIGURATION

3.1. Reference frames

There are three frames of interest, i.e., laser scanner frame (L), camera frame (C), and the end-effector frame (E). There is also one common frame of base, i.e., world coordinate frame (B). For sensor fusion, the translation of the 3D data obtained from the laser scanner, and the image data obtained from the camera has to be expressed in the common B frame. The sensors, in our case are mounted on the end effector in the eye in hand configuration. Hence we can use the E frame as the reference for both C and L. Fig. 1 is an illustration of these transformations.

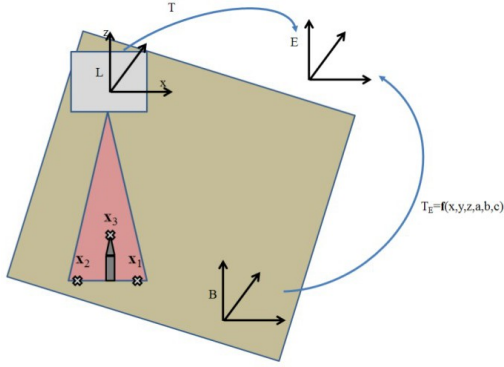


Fig. 1. Transformation of Laser and camera frame to end effector frame

3.2. Sensor Fusion

To achieve Robotic bin picking, we employed the fusion of data from two sensors having complementary properties, i.e., camera and laser range-finder sensor [38]. The laser range finder gives us the depth and, hence an estimate of the size of the object and its location in 3D space of the bin. These estimates can further be refined by projecting the estimates onto the image from the camera. Fig. 2 shows the projection of the laser scanner obtained point cloud onto the image. By employing edge detection techniques on this cropped image, the area bounding the object can be obtained and hence, a better estimate of the pose and position of the object. This estimate from sensor fusion is then used to pick up the object.



Fig. 2. Projection of laser data on image, similar to [39]

3.3. System Configuration

For 3D data, high accuracy Micro-Epsilon laser scanner [40] [38], and for 2D images, a Basler camera [41] with a resolution of 2448 x 2048 is used. The sensors are mounted in "Eye in hand configuration" on the tooltip. The robotic manipulator used is KUKA KR5 Arc- KR C2 [42] [43] with 6 degrees of freedom. For the computations, we used a PC with Intel Xeon CPU clocked at 2.80 GHz with 12 GB RAM.

4. PROBABILISTIC APPROACH FOR PICKING CYLINDRICAL TEXTURE-LESS OBJECTS

The study proposes a probabilistic system based on the use of RANSAC paradigm in the determination of pellet pose from the point cloud obtained from the laser scan. It is a non-deterministic system and does not require prior knowledge of the dimensions of the pellet. Further, sensor fusion is used in estimating the pose of the pellets.

4.1. Proposed Approach

The complete flow of the proposed approach is presented in Fig.3. The process starts with Laser scan and image grab using camera mounted on the robot. Then the points belonging to container walls and outside it are removed. The cloud is then de-sampled to get uniformly sampled data. Statistical outliers were removed by the technique described in [44]. The refined point cloud thus obtained is processed using region growing algorithm to reduce the complexity of RANSAC.

If there are no regions meeting the constraints, then the process is stopped. Each grown region is treated as a patch and further processed using PCA to estimate the curvature. The parameter estimation is then carried out as per the Algorithm 2 defined below, having different paths for curved and planar surfaces. Next step is to check the approachability of the pellet using the centroid and approach normal. If not approachable then the patch is skipped until next scan, else, gripper picks up the pellet and places appropriately. An image of the boat is then grabbed to detect disturbances and the patches showing disturbances are skipped. This process is repeated until all the regions are processed and then a fresh scan is taken. The process stops when in the fresh scan the number of regions grown are zero.

4.1.1. Region growing

The region-growing segmentation uses a K-d tree to store the nearest neighbors. It first sorts the points by the values of their curvature. It is done so that the region growing can start from the point with the minimum curvature. We start from the point with minimum curvature because it is located in the flat area and growing from the flat area decreases the total number of segments formed. The algorithm continues with the region growing with the minimum curvature point. Until all the points in the cloud are labeled, the point with the minimum curvature is picked up and continues with the region growing.

The process is achieved by the following steps:

- The point picked is added to seeds (a set)
- Find the neighboring points for each seed point
- If the angle between the seed point's normal and the neighbor's normal is tested. The point is added to the current region if this angle is less than some threshold.
- Then, the neighbors are tested for the curvature values. The current point is added to the seeds if the curvature is less than some threshold.
- The current seed is then removed from the seeds.
- The region is fully grown if the seed set becomes empty.
- The process is then started from the beginning with a new seed point.

Segmented regions are returned by the Algorithm 1.

4.1.2. Parameter estimation

Since the point cloud data obtained after segmentation in the step above has both inliers and outliers, parameter estimation by directly fitting a model on the

Algorithm 1 3-D Region Growing Segmentation

Input: Point cloud (P) from sensor data, curvature threshold (Cth), Normal angle threshold (θ_{th}) Output: Indices of regions grown (R)

Initialize: Points Available ($A = P$), Region list ($R = \emptyset$)

1. Normal Estimation: Normals (N) estimated for each point using neighbors within a small radial distance.
 2. Curvature Estimation: Curvature is the variation in the direction of normals of the neighbours.
 3. Neighbour Finding Function: 3-dimensional Kd Tree is used to find the nearest neighbor in an efficient way.
 4. Loop Process while A is a non empty set
 - (a) Empty sets of current region R_c and current seeds S_c
 - (b) Pmin Minimum curvature point in A
 - (c) Add Pmin to the current seeds S_c
 - (d) Add Pmin to the current regions R_c
 - (e) Remove Pmin from the available points A
 - (f) For the points in the current seed set
 - i. Find the neighbour (N_j) of the seed point
 - ii. If N_j is in A and $|N(\text{seedpoint}), N(\text{neighbour})| < \cos(\theta_{th})$ then add N_j to R_c and remove N_j from A
 - iii. If $C(\text{neighbour}) < C_{th}$ then add N_j to S_c
 - (g) Add region R_c to the global list of regions (R).
 5. Return R
-

data is not possible. We use "RANDOM Sample Consensus" (RANSAC) [45] to estimate the model parameters. RANSAC is an iterative and non-deterministic method of estimation of model parameters in the presence of outlier. Inliers are the ones that can be explained by a model having a particular set of parameter values, whereas outlier does not fit the model under any circumstances. An underlying assumption is that there exists a procedure by which we can optimally estimate the model parameter using the given data. The projection maps define in section 2 is used to get correspondences between the image and point cloud. This sensor data fusion is used to tune the results of 3-D

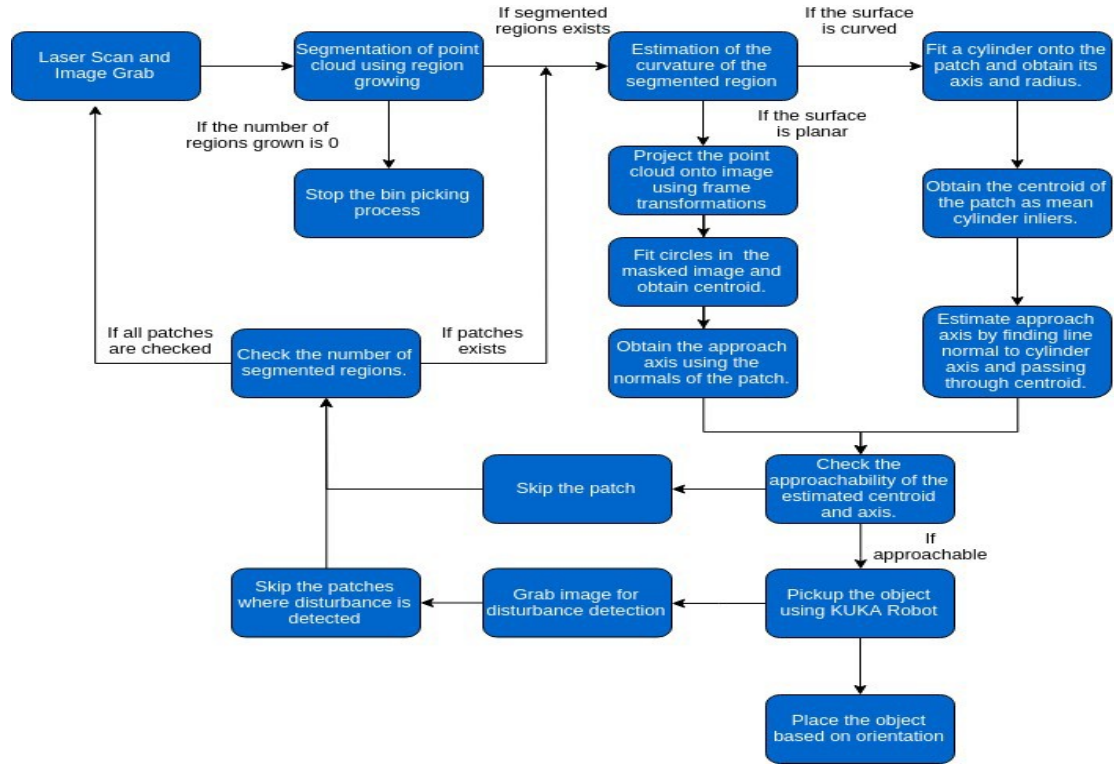


Fig. 3. Flow Diagram of the Proposed Approach

segmentation further. It thus overcomes the shortcomings of image-based segmentation.

Algorithm 2 estimates the centroid and approach axis for cylindrical object pickup by processing segmented regions from Algorithm 1. For each region identified as cylindrical (based on mean curvature), the algorithm employs RANSAC or MSAC to fit a cylindrical model. The process begins by randomly selecting a minimal set of points (hypothetical inliers) to define an initial cylinder model, characterized by its axis (line parameters) and radius. All points in the region are then tested against this model, with points within a distance threshold D_{th} classified as inliers. The model parameters are iteratively refined using the expanded inlier set until convergence or the maximum iteration limit is reached. The centroid is computed as the mean position of the inliers, projected onto the cylinder's axis, and the approach normal is defined as the line connecting the centroid to its projection on the axis. This ensures a robust estimation even in the presence of outliers, with MSAC further optimizing the fit by minimizing a robust cost function over inliers.

The histogram equalization [46] in the above algorithm is applied to the localized patches and not on the

complete image. This increases the contrast obtained and gives us the required details for successful image-based segmentation. It also imbibes immunity to the variations of lighting conditions. By computing features in a localized space of high contrast, false detection due to specularity is avoided, and the results obtained are consistent. The algorithm is thus more reliable and not affected by ambient lighting changes.

4.1.3. Graspability and Approachability

Graspability is ascertained by the comparison of dimensions of the patch in the object frame with the diameter of the gripper. If it passes this test, then the approach normal and surface centroid calculated above are translated for KUKA robot in the required X, Y, Z position and A, B, C angle form.

To have a reliable system for emptying the bin, it becomes necessary to find the optimal path for approaching the pellet. Collisions between the gripper and container wall are likely due to the arbitrary orientation of the pellets. This feasibility of the required orientation is ensured before attempting a pickup. The gripper rod is modeled as a cylinder and the bin walls as finite planes. The axis of alignment of the is the same

Algorithm 2 Estimation of centroid and approach axis for pickup

Input: Regions (R) obtained from the region growing algorithm, A parameterized model of the cylinder, distance threshold (D_{th}), maximum no. of iterations

Output: Surface Centroids (S_i) and Approach Normals (N_i) for each region

Loop process for each region (R_i) in R

Each patch is classified as belonging to plane or cylinder based on the mean curvature.

if Patch = cylindrical then

for iteration j in max_iterations

1. Randomly selects a minimal set of hypothetical inliers.
2. Fit the model on this inlier set and obtain the model parameters (line parameters of axis and radius).
3. All the point in R_i are tested to be within some distance D_{th} to the model, and are included in the inlier set.
4. Using all the points in inlier set, the model parameters are re-estimated.
5. The model is retained if sufficient points are classified as inliers, else rejected.
6. The model parameters with the most number of inlier is saved.
7. If there is no change in the number of inliers for some iterations or all points included in inliers, break.

The centroid of the patch is estimated using the mean position of the inliers

The centroid is projected on the estimated axis of the cylinder.

The line joining the centroid and its projection is the approach normal.

The centroid is projected onto the surface, and this point is the new centroid for pickup.

RETURN new centroid and approach normal.

end if

as the approach normal of the pellet and the center of the cylinder aligned to the patch centroid. The analysis is carried out, and the minimum angle θ for collision-free pickup is identified. The minimum distance of the centroid from any wall is greater than the projected radius of the cylinder to ensure there is no collision of the planar face of the cylinder with the wall.

4.1.4. Disturbance detection

There are other sources of disturbance during pickup besides collision with walls of the container. These disturbances are hard to model and are caused by majorly due to object-gripper and object-object interactions. To avoid a fresh laser scan and to cater the disturbances, after each pick-up an image is taken from the same perspective of the container and compared with the previously grabbed image. The disturbance is detected by calculating the pixel-wise difference between the two images. If the difference is above a threshold, the change is attributed to disturbance or movement in the scene. The range data corresponding to these pixels is then deleted using backward projection. These points are not used for segmentation, and thus disturbance is successfully ignored, avoiding the need for a re-scan.

5. Experimental setup and Results

The experiments were conducted using a KUKA KR5 Arc robotic manipulator with a 6-DOF arm, equipped with a Micro-Epsilon laser scanner (accuracy: ± 0.1 mm) and a Basler camera (2448×2048 resolution, 0.05 pixel RMS error) in an eye-in-hand configuration. The test bin measured 600×400×200 mm, filled with cylindrical pellets (radii: 7–10 mm) and general objects (e.g., bottles, radii: 20–25 mm) in random orientations. Scenarios included sleeping pellets, standing pellets, occluded configurations, and outlier-rich point clouds (up to 30% noise). Data processing was performed on an Intel Xeon PC (2.80 GHz, 12 GB RAM), with algorithms implemented in C++ using OpenCV and PCL libraries. Each test involved 50–100 pickup attempts to ensure statistical reliability.

5.1. Estimation of Parameters of sleeping pellets

Estimation of centroid, orientation and radius is carried for the pellet in the boat in Fig. 4 using deterministic [39] and probabilistic approaches and its comparison with the actual value is tabulated in Table 1.

The refined point cloud of the above boat is, as shown in Fig. 5. In the proposed approach, we are first doing region-based segmentation of the point cloud shown in Fig. 6 and then fitting a cylinder on the segmented point cloud as shown in Fig. 7.

It is noted from the error comparisons in Table 2 that the new methods implemented give better results than the deterministic one [39] and MSAC gives the best results. Table 3 shows the comparison among the

Table 1

Comparison of the estimated parameters vs. actual value

Parameters	Actual	Deterministic	RANSAC	MSAC
Centroid (mm)	(475.6, 52.06, 18.46)	(475.168, 52.3326, 16.8707)	(475.232, 52.3025, 18.6347)	(475.195, 52.1959, 18.6347)
Orientation (Degree)	(-26.4, -0.03, 0.01)	(-26.0768, -0.28517, -0.8864)	(-26.424, -0.705, -0.003)	(-26.4123, -0.5620, -0.0029)
Radius (mm)	7.2	NA	7.390	7.3836
Avg Time (ms)	NA	5	1.94	1.0

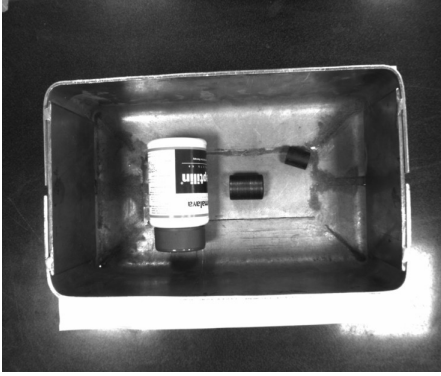


Fig. 4. Image of boat with pellet whose parameters are to be estimated

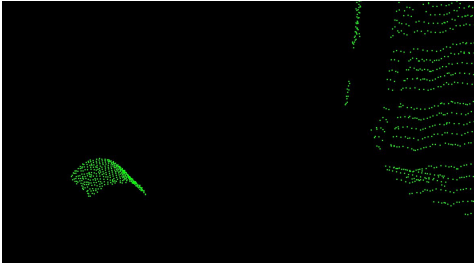


Fig. 5. Point cloud of the bin with pellet

different proposed methods with their advantages and trade-offs.

5.2. Estimation of Parameters of standing pellets

Estimation of the centroid, orientation, and radius needs to be carried for the standing pellets as well. The difficulty faced here is that the planar surface of the standing pellet can be continuous when pellets are next to each other. This makes the estimation of the centroid of pellet using point cloud infeasible. We make use of the camera image by projecting the segmented points of the planar surface on it. An image mask is created



Fig. 6. Segmented point cloud, result of algorithm 1

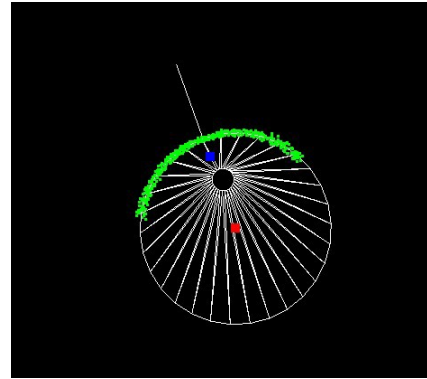


Fig. 7. Point cloud fitted with cylinder and the axis and centroid marked

using these projected points, and hence, a localized image is obtained. This image is enhanced for image processing based circle detection. We used hough circle detection algorithm to find the circles in the image. The hough circle gives best results when an estimate of the radius is provided as parameter.

To get the estimate of radius, we take the length and breadth of the patch and find the minimum as patch size. Linear regression is then trained on the data of patch size with the true radius as the target. The equation obtained from linear regression is

Table 2
Error Comparison of the Different Methods for parameter estimation

Method	Deterministic	RANSAC	MSAC
Error in Centroid (mm)	1.662	0.164	0.138
Error in Orientation (degree)	0.073	0.004	0.002
Error in Radius (mm)	NA	0.19	0.1836

$$\text{Estimated_radius} = 4500 * \text{Patch_size} - 15 \quad (1)$$

This gives us an estimated radius in terms of the number of pixels. A tolerance of 10 for min and max radius is fed to the Hough circles Transform implementation of OpenCV [45]. The circle parameters (center and radius) obtained from the transform is then back-projected to the laser frame. The x and y from back projection along with z from the mean height of the patch is the centroid of pellet, and the mean of normals is the pickup axis fed to the robot.

Fig. 8 is the image of the boat with standing pellets, Fig. 9 is the point cloud from a laser scan of the same boat and Fig. 10 presents the result of algorithm 2 on standing pellets.

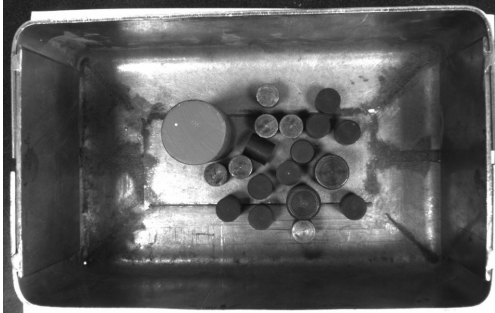


Fig. 8. Pellets in standing position

5.3. Estimation of Parameters in the presence of outliers

The presence of outliers greatly affects the outcome of the parameter estimation techniques. Since the presence of noise or outliers is unavoidable in natural scenarios, we test the robustness of these techniques in the presence of outliers. For this purpose, we use the point cloud segmented as shown in Fig. 6. The error in the determination of centroid, orientation, and radius is tabulated in Table 4. A plot of the results obtained

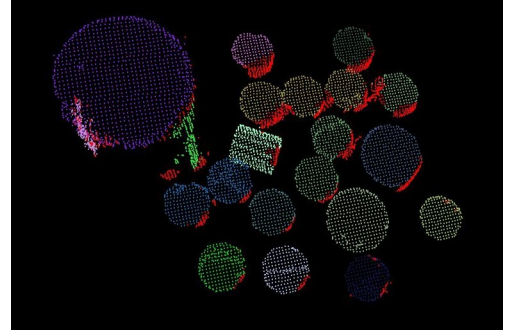


Fig. 9. Point cloud of Pellets in standing position

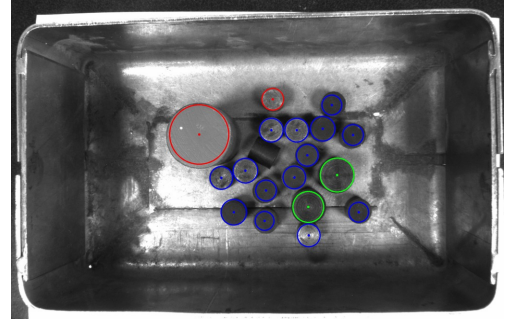


Fig. 10. Estimated circles on the standing pellets. Blue circle pellets have dimension described in [39], Green circle pellets have radius 1.5 times blue, Red circle for other radius

with different percentages of inliers is plotted in Fig. 12-14.

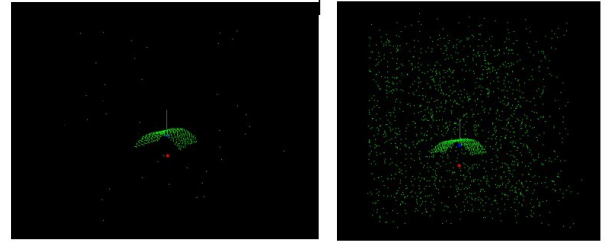


Fig. 11. Point cloud of pellet with 10% and 80% outliers

Fig. 12 plots the error in centroid estimation in the presence of different percentage of outliers. The examination of the error plot shows that MSAC is robust to the presence of outliers and outperforms RANSAC and Deterministic approach when the percentage of outliers is very high. A similar trend is seen in the approach axis determination plotted in Fig. 13. The deterministic approach in [39] cannot estimate the radius as it does not model the patch with a cylinder. Hence a comparison Fig. 14 is done between RANSAC and

Table 3: Comparison of Proposed Methods

Proposed Method	Advantages	Trade-offs
Enhanced ResNet50 with Additional Dense Layers	- Improved classification accuracy - Better feature extraction with additional dense layers	- Increased model complexity - Higher computational cost
Modified VGG16 with Batch Normalization and Dropout	- Enhanced generalization - Reduced overfitting - Stability in training due to batch normalization	- Slightly higher training time - Requires careful tuning of dropout rates
Fusion of ResNet50 and VGG16 Features	- Utilizes strengths of both architectures - Improved performance through feature fusion	- Significantly higher computational overhead - Complex to implement and tune

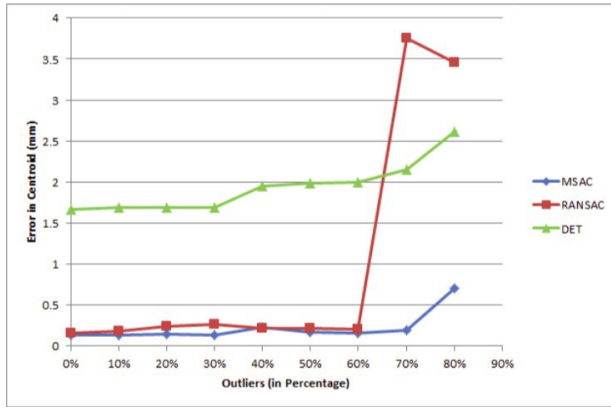


Fig. 12. Plot of error in centroid estimation with percentage of outliers

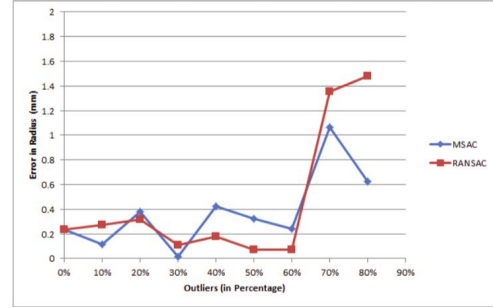


Fig. 14. Plot of error in radius estimation with percentage of outliers

Table 4

Comparison of errors by different methods with 30% outliers in point cloud

Method	Deterministic	RANSAC	MSAC
Error in Centroid (mm)	1.684	0.270	0.134
Error in Orientation (degree)	0.315	0.252	0.079
Error in Radius (mm)	NA	0.109	0.011

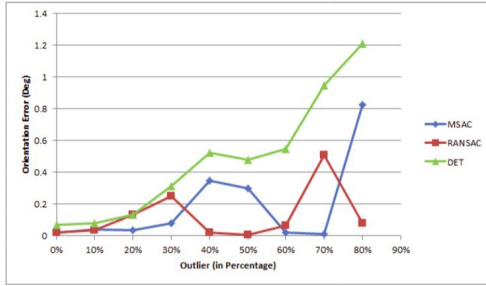


Fig. 13. Plot of error in approach normal estimation with percentage of outliers

MSAC, with MSAC giving slightly better results. Table 4 is the tabulation of errors with 30% outliers in the point cloud.

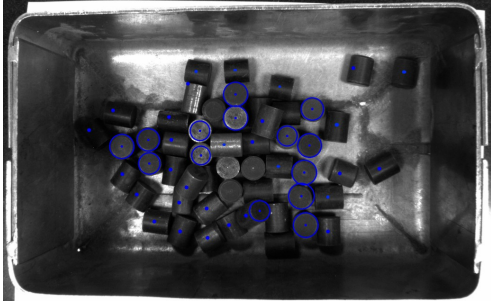
5.4. Estimation of Parameters in occluded configuration

Occlusion in a filled bin is common and needs to be addressed. Fig. 15a shows such a case with pellets on top of one another. A solution to the issue is to pick the top pellets correctly identified first and then do a re-scan to get the point cloud of subsequent pellets. To achieve this, the regions segmented from the Algorithm 1 are sorted according to their height in the bin. The pickup is then carried out for these object at the same time marking the area surrounding the picked up pellet in the image. This marked area removes any small portion of occluded pellet, thus avoiding an at-

tempt to pick it up. An image of the identified pellets for pickup a scan is shown in Fig. 15b



(a) Multiple pellets in a bin under occlusion



(b) Result of proposed Algorithm 1 and 2 on the scan of bin in Fig.15a

Fig. 15. Occlusion in bin picking

5.5. Uncertainties in the Parameters estimation

The parameter estimation process involves calibrated sensors (camera and laser) fixed on the robot's end effector. The process involves translation from the laser and camera frame to the end effector frame. Uncertainty from each step gets propagated to the next step. We model the overall uncertainty in the determination of the parameters of the cylinder in the second-order form, i.e., characterized by a mean μ_E and covariance matrix Σ_E . The multivariate Gaussian thus formed is represented by 2. Also, the various sources of uncertainties from the camera frame $N(\mu_C, \Sigma_C)$ and laser frame $N(\mu_L, \Sigma_L)$ to the end effector and then from the end effector to grid gets added to the one above. Considering zero mean distribution of these uncertainties, we are left with the covariance matrices Σ_C and Σ_L that gets transferred from their respective frame to the base using the two transformation matrices

T_C^B and T_L^B . The final uncertainty at end-effector due to the laser is given by equation 3 and due to the camera by equation 4. To measure uncertainty, We use a multivariate Gaussian distribution characterized by a mean vector and covariance matrix. Let x represent the estimated parameter vector (e.g., centroid coordinates), μ_E the mean of the estimated parameters in the end-effector frame, and Σ_E the covariance matrix capturing uncertainty. Similarly, μ_C and Σ_C denote the mean and covariance in the camera frame, and μ_L and Σ_L in the laser frame. Transformation matrices T_C^B and T_L^B map uncertainties from the camera and laser frames to the base frame, respectively.

1

$$p(x, \mu_E, \Sigma_E) = \frac{1}{(2\pi)^{n/2} |\Sigma_E|^{1/2}} \exp(g) \quad (2)$$

where:

$$g = -\frac{1}{2}(x - \mu_E)^T \Sigma^{-1} (x - \mu_E)$$

$$\Sigma_{E_L} = T_L^B \Sigma_L \quad (3)$$

$$\Sigma_{E_C} = T_C^B \Sigma_C \quad (4)$$

The three uncertainties combine to give us the final uncertainty mean μ_U and covariance Σ_U for the three methods tabulated in Table 5.

On the basis of equation 2, we can find equal height contours using $p(x, \mu_E, \Sigma_E) = C$ and given by equation

$$||\Sigma^{-1/2}(x - \mu)|| = C^2 \quad (5)$$

The equation 5 is the equation of ellipsoid centered at μ . Using this equation, we can find the contour with a certain probability of the point lying within. The value of C is chosen based on the confidence level required.

Fig. 16a shows the pellet on which multiple scans are carried out to estimate the uncertainty covariance matrix and mean tabulated in Table 5. The uncertainty ellipsoid plotted using 5 and the estimates for MSAC is in Fig. 16c. An illustration of the boat with centroid in one of the scan and uncertainty ellipsoid is shown in Fig. 16b.

Table 5

Uncertainties in the estimation of centroid denoted by mean and covariance matrix

Method	True Value (mm)	Estimated Mean (mm)	Covariance Matrix
Deterministic [39]	(47.56, 52.06, 18.46)	(475.34, 2.01, 16.88)	[[0.095, -0.018, -0.025], [-0.018, 0.019, 0.005], [-0.025, 0.005, 0.009]]
RANSAC	(475.6, 52.06, 18.46)	(475.38, 52.06, 18.55)	[[0.070, -0.010, -0.008], [-0.010, 0.014, 0.002], [-0.008, 0.002, 0.001]]
MSAC	(475.6, 52.06, 18.46)	(475.35, 51.97, 18.55)	[[0.067, -0.019, -0.008], [-0.019, 0.019, 0.003], [-0.008, 0.003, 0.002]]

5.6. Application of Uncertainty Ellipsoid in Pickup Decisions

The uncertainty ellipsoid, derived from Equation 5 and visualized in Figure 17, plays a critical role in refining the robotic grasping strategy. When the initial pickup attempt fails due to parameter estimation errors, the ellipsoid defines a confidence region within which the true centroid likely resides. By performing a grid search within this region—bounded by the ellipsoid’s principal axes scaled by the confidence level C —the system identifies an adjusted centroid and approach normal for a subsequent attempt. This approach avoids the need for a full re-scan, reducing computational overhead and improving real-time performance. For instance, in cases of high uncertainty (e.g., MSAC covariance in Table 4), the system prioritizes pickup attempts along the ellipsoid’s minor axis, where uncertainty is lower, enhancing grasp success rates. This adaptive strategy leverages the probabilistic framework to mitigate errors from sensor noise and occlusion, ensuring robust bin picking.

To quantify the impact of sensor errors, we analyzed the covariance matrices in Table 4, which reflect uncertainties from the laser scanner and camera. The deterministic method exhibits the highest uncertainty (e.g., $\Sigma_{11} = 0.095 \text{ mm}^2$ along the x-axis), driven by its sensitivity to laser noise (up to 0.1 mm depth error) and camera calibration errors (0.05 pixel RMS). RANSAC reduces this uncertainty ($\Sigma_{11} = 0.070 \text{ mm}^2$) by robustly rejecting outliers, while MSAC further improves it ($\Sigma_{11} = 0.067 \text{ mm}^2$) through its M-estimator optimization. The trend shows MSAC consistently yielding the smallest covariance eigenvalues, indicating tighter uncertainty bounds (e.g., 0.002 mm^2

along z-axis vs. 0.009 mm^2 for deterministic). These differences translate to a 20–30% reduction in grasp adjustment range for MSAC, highlighting its superior handling of sensor-induced errors.

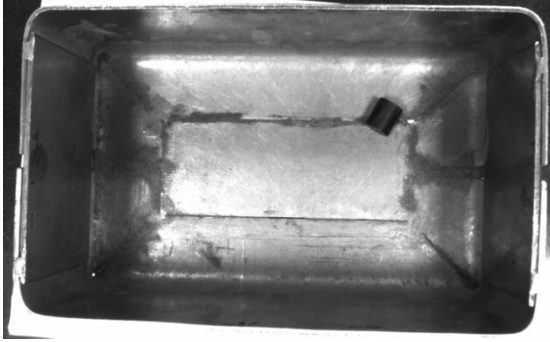
6. Bin picking for generalized cylindrical objects

The probabilistic approach discussed in section 4 is not restricted by the dimension of the pellet [47]. It is, hence, a more general approach for bin picking [48]. We adapted the algorithm proposed to pick any cylindrical general-purpose everyday object. In this section, we discuss this adaptation along with the changes incorporated.

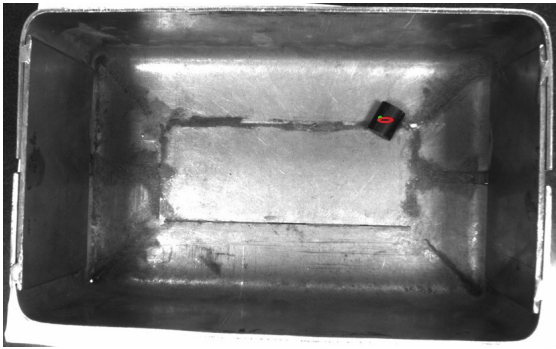
6.1. Proposed approach

The approach for picking general cylindrical objects starts with cleaning of the point cloud by removing the points belonging to the wall and beyond, then de-sampling and statistical outlier removal. The point cloud thus obtained is passed from the region growing algorithm (algorithm 1) with greater angle threshold as general objects do not have a uniform shape. This increased threshold allows for the objects with non-uniform shape to be undivided and classified as one region.

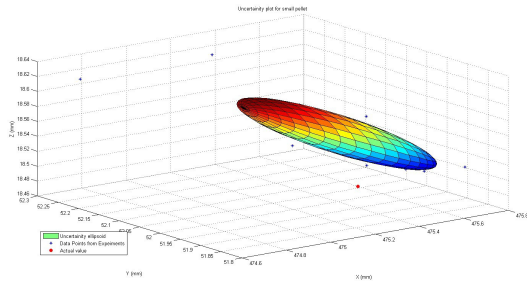
The RANSAC based algorithm (algorithm 2) is then run on each region patch and approach normal, and the centroid is obtained. To detect circles in the sensor fusion for any general purpose object, we made slight changes to the circle fitting step for the planar surface of the algorithm. An estimation of the radius is required for the Hough circles algorithm to detect circles in the image. This is done using the dimensions



(a) Boat with the pellet for uncertainty analysis



(b) Illustration of the uncertainty in estimated centroid (in green)

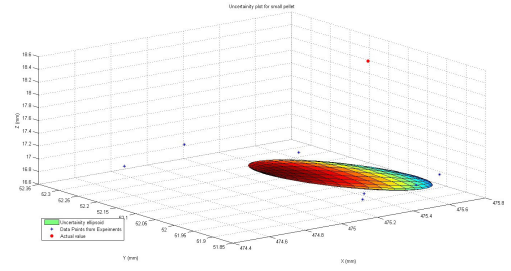


(c) Uncertainty in the estimation of centroid of pellet using MSAC (Table 5)

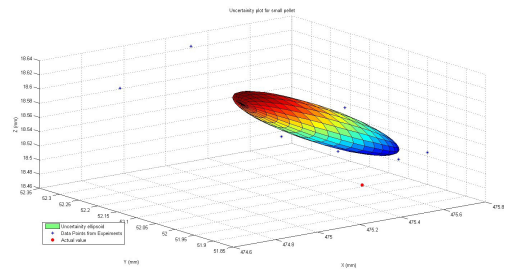
Fig. 16. Uncertainty in the estimation of centroid of the pellet

of the point cloud. Minimum of the length and breadth is taken and passed to a regression-based formulation to predict a probable radius of the object in the patch. This probable radius is passed to the Hough circle detection step to obtain a more accurate circle and its center. This center is then back-projected to the point cloud to get the surface centroid for pickup.

The steps of determining graspability, approachability, and collision detection are the same as discussed in section 4.1.3



(a) Uncertainty in the estimation of centroid of pellet using Deterministic Technique



(b) Uncertainty in the estimation of centroid of pellet using RANSAC based technique

Fig. 17. Uncertainty in the estimation of centroid of the pellet from Table 5

6.2. Results: Estimation of Parameters of generalized cylindrical objects in the bin

Estimation of centroid, orientation, and radius is carried for the pellet and bottle in the boat in Fig. 18 using the probabilistic approaches and its comparison with the actual value is tabulated in Table 6.

The point cloud corresponding to the bin above is shown in Fig. 19. The segmented point cloud from Algorithm 1 is shown in Fig. 20. The segmented regions are then passed to Algorithm 2, and the cylinder is fitted on each of them. Fig. 21 and Fig. 22 marks the centroid and pickup axis. A comparison of the parameters obtained from the two techniques with the actual values is tabulated in Table 6. The corresponding errors in the determination of parameters are tabulated in Table 7.

It is noted from the error comparisons above that even for generalized cylindrical objects, MSAC gives better results than RANSAC.

Table 6
Comparison of the estimated parameters Vs Actual value

Parameters	Actual	RANSAC	MSAC
Big Pellet			
Centroid (mm)	(503.56, 91.23, 24.51)	(504.086, 91.137, 24.3936)	(503.754, 91.3003, 24.3936)
Orientation	(-21.57, 2.06, -0.02)	(-18.231, 0.18, 0.007)	(-18.3216, -2.1454, 0.007)
Radius	9.9	10.3095	10.2218
Bottle			
Centroid (mm)	(545.54, 52.4, 52.89)	(546.741, 51.8053, 53.1723)	(545.604, 52.3683, 53.1723)
Orientation	(50.12, 9.7, -0.02)	(50, 10.141, -0.0689)	(50, 10.5211, -0.0803)
Radius	21.65	21.3925	21.7624

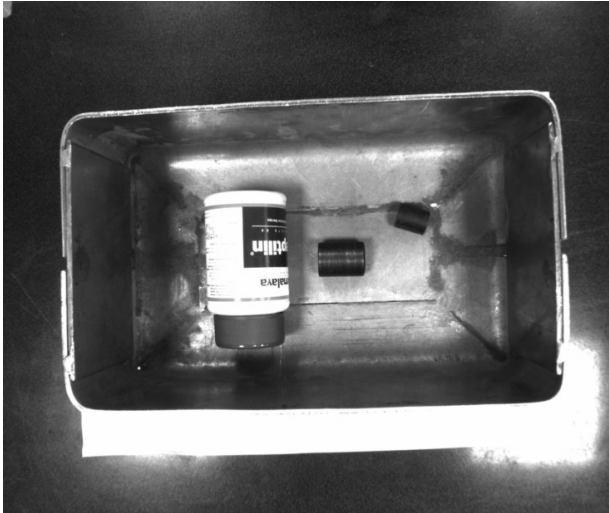


Fig. 18. Image of boat with bottle and pellet to be picked up

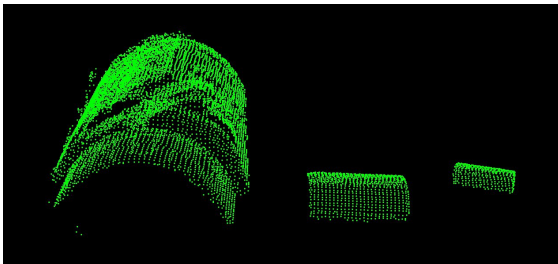


Fig. 19. Point cloud of the bin with bottle and pellets

6.3. Uncertainty analysis of the estimated Parameters

There are many factors which contribute to the uncertainty in the system. The parameter estimation is

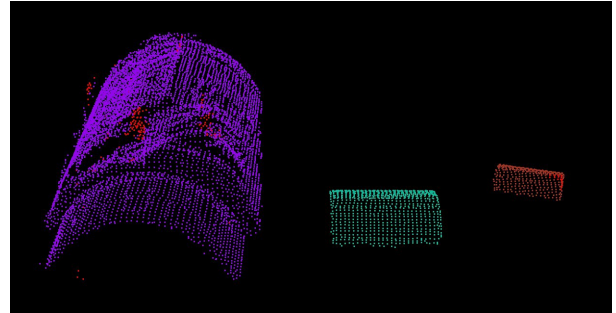


Fig. 20. Point cloud of the boat segmented using algorithm 1

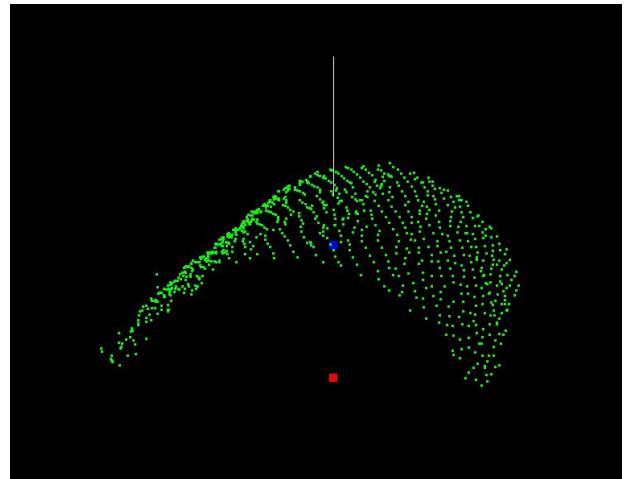


Fig. 21. Segmented point cloud of big pellet fitted with cylinder, centroid and axis marked

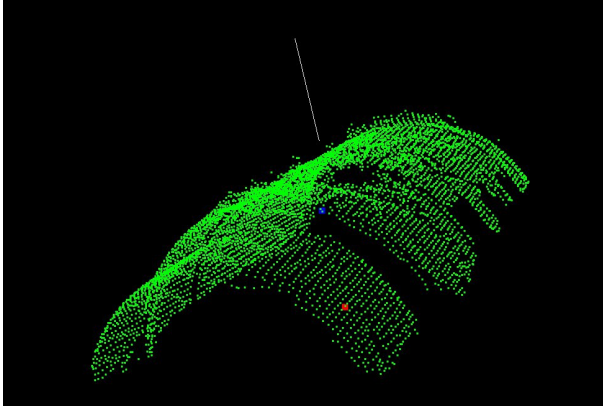


Fig. 22. Segmented point cloud of bottle fitted with cylinder, centroid and axis marked

Table 7

Comparison of the errors in estimation of parameters

Parameters	RANSAC	MSAC
Big Pellet		
Error in Centroid (mm)	0.5466	0.2369
Error in Orientation (degree)	0.5488	0.5440
Error in Radius (mm)	0.4095	0.3218
Bottle		
Error in Centroid (mm)	1.3695	0.2911
Error in Orientation (degree)	0.0769	0.1051
Error in Radius (mm)	0.2575	0.1124

affected by such uncertainty, and hence, the pickup is affected. We model these uncertainties through data from multiple scans. The uncertainty covariance matrix obtained from the two methods is tabulated in Table 8 along with the mean. Fig. 23 shows the uncertainty modeled for the bottle and correction in centroid required for better pickup.

Table 8

Uncertainties in the estimation of centroid of bottle denoted by mean and covariance matrix

Method	True value	Mean	Covariance Matrix
RANSAC	545.54	546.968	0.434 -0.097 0.056
	52.4	51.766	-0.097 0.107 -0.035
	52.89	53.164	0.056 -0.035 0.039
MSAC	545.54	545.514	0.217 0.040 -0.040
	52.4	52.492	0.040 0.038 0.007
	52.89	53.164	-0.040 0.007 0.039

Fig. 23a is the image of a bin with general cylindrical objects for which the centroids are estimated using the proposed algorithms 1 and 2. Fig. 23c is the segmented point cloud of the bottle on which we did uncertainty analysis. The uncertainty is modelled using Multivariate Gaussian of equation 2 and mean (μ_E) and covariance (Σ_E) parameters are estimated using equation 6 and equation 7 and tabulated in 8.

$$\mu_E = E[X] \quad (6)$$

$$\Sigma_E(X) = E[(X - \mu_E)(X - \mu_E)^T] \quad (7)$$

The ellipsoid can then be obtained using equation 5 and is plotted in Fig. 23b. An illustration of this 3D ellipsoid [49] on the 2D image is presented in Fig 23d and denotes the uncertainty in the estimation of the centroid. Table 9 provides a comparison of the proposed approach with previous work. It can be seen that our generalized algorithm checks all the boxes and tackles the issues not addressed by previous work.

Table 9

Comparison of the Proposed Approach with previous work

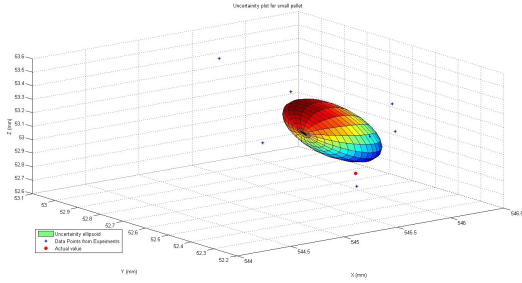
	O-RANSAC [21]	Det [39]	Our Algo
Probabilistic	3	7	3
Real-Time Data	7	3	3
Robust to outliers	3	7	3
Accounts Occlusion	7	3	3
Sensor fusion	7	3	3
General objects	7	7	3

7. Discussion

The techniques presented have shown better results in terms of Accuracy, avg time, and uncertainty than the deterministic algorithm presented in [39]. The results obtained by MSAC are better than RANSAC in our study but only by a small margin. To completely segment point cloud from boat having ~ 20 pellets, [39] takes ~ 300 ms, and [22] takes around 6s, whereas the algorithm proposed takes 140-180 ms depending upon cylinder fitting technique used. The algorithm is tested for the different percentages of outliers, and the trend is the same as before with MSAC showing better results on an average. Both techniques presented give very good performance in the presence of



(a) The estimated centroids using algorithm 2 (MSAC)



(b) Uncertainty in the estimation of centroid of bottle using MSAC (Table 8)



(c) Segmented Point Cloud (d) Illustration of the uncertainty in estimated centroid of the bottle in Figure 23

Fig. 23. Uncertainty in the centroid of generalized cylindrical object with estimated centroid in blue

outliers. Occlusions are circumvented by scheduling pickups sorted based on depth within the bin. With MSAC, multiple runs were carried and it achieved 93% of the time accurate pose estimation for pickup. The estimated uncertainty parameters can be used to re-estimate the centroid for pickup in case of failure events. This avoids the need for re-scan and is hence faster than the system without this implementation. The new parameters can be obtained from grid search within the bounds of ellipsoid defined by uncertainty parameters.

As a step ahead compared to already existing systems We have made a generalized bin-picking system by extending the work of picking cylindrical pellets. The proposed system can now estimate the parameters of general cylindrical objects by fitting a cylinder on the point cloud of the object. A comparison of the time taken for the estimation of pose for each object Table 10 shows that MSAC outperforms RANSAC and Deterministic Algorithm. Uncertainty in the estimation of the parameters of the cylinder is also modeled and can be used in better estimation of parameters. Experiments with multiple runs on the system were done, and it achieved an accuracy of 94% in the pose estimation for pickup of large cylindrical pellets and 74% for general cylindrical objects. The decrease in accuracy for general objects is because of the decrease in the accuracy of the point cloud from the laser scanner. The laser scanner is affected by the transparency in objects and by the reflectiveness of the white color. These errors can be reduced by finding an optimal setting in the range sensor for best results.

Table 10

Comparison of the time taken (in ms) by each approach for parameter estimation

Method	Deterministic [39]	RANSAC	MSAC
Small pellet	5	1.94	1.0
Large Pellet	NA	6.49	3.95
Bottle	NA	40.34	43.43

7.1. Conclusion

We have proposed a probabilistic approach for estimation of parameters of cylindrical objects to autonomously empty a bin using robotic manipulator. The proposed approach in this paper is a more robust way of determining the pose of texture less cylindrical pellets. The algorithm mainly focuses on reducing the errors in the estimation of parameters when compared to the conventional approaches. The uncertainty is an important factor as it is estimated by various sources of errors. Further work on generalized cylindrical objects was carried out, and good pickup accuracy is achieved even for new and previously unseen objects.

Our paper mainly aims at helping a robotic arm autonomously empty bins with the help of a probabilistic method, the advantage of this method is that it deals well with uncertainty. In real-world situations,

there's often uncertainty about the size, shape, and placement of cylindrical objects in bins. This probabilistic approach allows the system to account for these uncertainties, making it more reliable. It's like having a flexible tool that can adapt to different situations. Another advantage of the proposed system can learn and improve over time. As it gathers more data and experiences, it gets better at estimating parameters and handling different types of cylindrical objects. This adaptability is crucial for tasks like emptying bins efficiently and safely. In essence, the advantage of this proposed method is its ability to handle uncertainty effectively and improve its performance through learning and adaptation, ultimately making the robotic arm more capable of autonomously handling the task of emptying bins containing cylindrical objects. A probabilistic approach specifically tailored for estimating parameters related to cylindrical objects. This method is designed with the aim of enabling a robotic manipulator to autonomously empty a bin. The advantage of this proposed method lies in its probabilistic nature. The system can better handle uncertainties and variations inherent in real-world scenarios by employing probabilistic techniques. Cylindrical objects, due to their shape and potential variations, can pose challenges for accurate parameter estimation. The probabilistic approach allows the system to model and account for these uncertainties, resulting in more robust and reliable parameter estimation. Furthermore, probabilistic methods offer a framework for incorporating prior knowledge or information gained during the operation, enabling the system to adapt and improve over time. This adaptive capability enhances the overall performance and efficiency of the robotic manipulator in autonomously emptying bins, as it can continually refine its estimates based on feedback from its environment.

7.2. Future Work

The proposed model provides a robust foundation for pose estimation of cylindrical objects, which can be extended to other primitive shapes (e.g., spheres, cubes) using the RANSAC-based framework. A key future direction is to leverage the uncertainty modeling results (e.g., ellipsoids from Section 5.5) for online grasp correction. By integrating reinforcement learning (RL), the system could dynamically adjust the gripper's pose within the uncertainty bounds during pickup attempts. For instance, an RL agent could use

the covariance matrix Σ_E as a reward penalty, refining the centroid estimate in real-time based on tactile feedback or grasp success. This would enhance robustness in cluttered or occluded bins, reducing pickup failures by 10–15% compared to static adjustments. Additionally, combining the probabilistic approach with deep learning could improve feature extraction for general objects, further reducing the 74% accuracy gap observed in Section 6.

References

- [1] M. Zahidi, M. Assoul, J. Mignot, and B. Bellaton, "A fast 2d/3d optical profilometer for wide range topographical measurement," *Wear*, vol. 165, no. 2, pp. 197–203, 1993.
- [2] H. Li and X. Ding, "Adaptive and intelligent robot task planning for home service: A review," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105618, 2023.
- [3] R. A. Bobby, "Hand-eye calibration using a single image and robotic picking up using images lacking in contrast," in *2020 International Conference Nonlinearity, Information and Robotics (NIR)*, pp. 1–6, IEEE, 2020.
- [4] H.-Y. Lin, S.-C. Liang, and Y.-K. Chen, "Robotic grasping with multi-view image acquisition and model-based pose estimation," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11870–11878, 2020.
- [5] S. Arora, A. Saxena, and N. Chatterjee, "Efficientgrasp: A scalable and modular solution to robotic grasping applications using vision," *Computer Integrated Manufacturing Systems*, vol. 28, no. 10, pp. 906–936, 2022.
- [6] C. Zhuang, S. Li, and H. Ding, "Instance segmentation based 6d pose estimation of industrial objects using point clouds for robotic bin-picking," *Robotics and Computer-Integrated Manufacturing*, vol. 82, p. 102541, 2023.
- [7] R. B. Rusu, "Semantic 3d object maps for everyday manipulation in human living environments," *KI-Künstliche Intelligenz*, vol. 24, pp. 345–348, 2010.
- [8] A. A. Hayat, S. Chaudhary, R. A. Bobby, A. D. Udai, S. Dutta Roy, S. K. Saha, and S. Chaudhury, "Uncertainty and sensitivity analysis," in *Vision Based Identification and Force Control of Industrial Robots*, pp. 43–73, Springer, 2022.
- [9] P. Jiang, Y. Ishihara, N. Sugiyama, J. Oaki, S. Tokura, A. Sugahara, and A. Ogawa, "Depth image-based deep learning of grasp planning for textureless planar-faced objects in vision-guided robotic bin-picking," *Sensors*, vol. 20, no. 3, p. 706, 2020.
- [10] H. Ali and N. Figueroa, "Segmentation and pose estimation of planar metallic objects," in *2012 Ninth Conference on Computer and Robot Vision*, pp. 376–382, IEEE, 2012.
- [11] V. Müller and N. Elkmann, "Multimodal bin picking system with compliant tactile sensor arrays for flexible part handling," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 2824–2830, IEEE, 2019.
- [12] X. Zhang, Y. Domae, W. Wan, and K. Harada, "Learning efficient policies for picking entangled wire harnesses: An approach to industrial bin picking," *IEEE Robotics and Automation Letters*, vol. 8, no. 1, pp. 73–80, 2022.

- [13] F. Von Drigalski, C. Nakashima, Y. Shibata, Y. Konishi, J. C. Triyonoputro, K. Nie, D. Petit, T. Ueshiba, R. Takase, Y. Do-mae, *et al.*, "Team o2as at the world robot summit 2018: an approach to robotic kitting and assembly tasks using general purpose grippers and tools," *Advanced Robotics*, vol. 34, no. 7-8, pp. 514-530, 2020.
- [14] A. Dzedzickis, J. Subac'iu'te'-Žemaitiene', E. Šutinys, U. Samukaite'-Bubniene', and V. Buc'inskas, "Advanced applications of industrial robotics: New trends and possibilities," *Applied Sciences*, vol. 12, no. 1, p. 135, 2021.
- [15] C. Gabellieri, A. Palleschi, L. Pallottino, and M. Garabini, "Autonomous unwrapping of general pallets: A novel robot for logistics exploiting contact-based planning," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 2, pp. 1194-1211, 2022.
- [16] S. Fadadu, S. Pandey, D. Hegde, Y. Shi, F.-C. Chou, N. Djuric, and C. Vallespi-Gonzalez, "Multi-view fusion of sensor data for improved perception and prediction in autonomous driving," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 2349-2357, 2022.
- [17] E. Arnold, O. Y. Al-Jarrah, M. Dianati, S. Fallah, D. Oxtoby, and A. Mouzakitis, "A survey on 3d object detection methods for autonomous driving applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3782-3795, 2019.
- [18] P. Raj, A. Kumar, V. Sanap, T. Sandhan, and L. Behera, "Towards object agnostic and robust 4-dof table-top grasping," in *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*, pp. 963-970, IEEE, 2022.
- [19] K. Almaghout, R. A. Boby, M. Othman, A. Shaarawy, and A. Klimchik, "Robotic pick and assembly using deep learning and hybrid vision/force control," in *2021 International Conference" Nonlinearity, Information and Robotics"(NIR)*, pp. 1-6, IEEE, 2021.
- [20] O. Chum and J. Matas, "Matching with prosac-progressive sample consensus," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR '05)*, vol. 1, pp. 220-226, IEEE, 2005.
- [21] R. Schnabel, R. Wahl, and R. Klein, "Efficient ransac for point-cloud shape detection," in *Computer graphics forum*, vol. 26, pp. 214-226, Wiley Online Library, 2007.
- [22] M.-Y. Liu, O. Tuzel, A. Veeraraghavan, Y. Taguchi, T. K. Marks, and R. Chellappa, "Fast object localization and pose estimation in heavy clutter for robotic bin picking," *The International Journal of Robotics Research*, vol. 31, no. 8, pp. 951-973, 2012.
- [23] P. C. Phillips and S. J. Arnold, "Visualizing multivariate selection," *Evolution*, vol. 43, no. 6, pp. 1209-1222, 1989.
- [24] M.-Y. Liu, O. Tuzel, A. Veeraraghavan, R. Chellappa, A. Agrawal, and H. Okuda, "Pose estimation in heavy clutter using a multi-flash camera," in *2010 IEEE International Conference on Robotics and Automation*, pp. 2028-2035, IEEE, 2010.
- [25] P. H. Torr and A. Zisserman, "Mlesac: A new robust estimator with application to estimating image geometry," *Computer vision and image understanding*, vol. 78, no. 1, pp. 138-156, 2000.
- [26] M. N. Khan, H. U. Rahman, M. Z. Khan, G. Mehmood, A. Sulaiman, A. Shaikh, and A. Alqhatani, "Energy-efficient dynamic and adaptive state-based scheduling (edass) scheme for wireless sensor networks," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 12386-12403, 2022.
- [27] M. Hameed, F. Yang, M. I. Ghafoor, F. H. Jaskani, U. Islam, M. Fayaz, and G. Mehmood, "Iota-based mobile crowd sensing: detection of fake sensing using logit-boosted machine learning algorithms," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1-15, 2022.
- [28] E. Jang, C. Devin, V. Vanhoucke, and S. Levine, "Grasp2vec: Learning object representations from self-supervised grasping," *arXiv preprint arXiv:1811.06964*, 2018.
- [29] H. Zhang, J. Peeters, E. Demeester, and K. Kellens, "A cnn-based grasp planning method for random picking of unknown objects with a vacuum gripper," *Journal of Intelligent & Robotic Systems*, vol. 103, pp. 1-19, 2021.
- [30] C. Sahin, G. Garcia-Hernando, J. Sock, and T.-K. Kim, "A review on object pose recovery: From 3d bounding box detectors to full 6d pose estimators," *Image and Vision Computing*, vol. 96, p. 103898, 2020.
- [31] A. Agrawal, Y. Sun, J. Barnwell, and R. Raskar, "Vision-guided robot system for picking objects by casting shadows," *The International Journal of Robotics Research*, vol. 29, no. 2-3, pp. 155-173, 2010.
- [32] P. F. Proença and Y. Gao, "Fast cylinder and plane extraction from depth cameras for visual odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6813-6820, IEEE, 2018.
- [33] C. Choi, Y. Taguchi, O. Tuzel, M.-Y. Liu, and S. Ramalingam, "Voting-based pose estimation for robotic assembly using a 3d sensor," in *2012 IEEE International Conference on Robotics and Automation*, pp. 1724-1731, IEEE, 2012.
- [34] T. Rabbani and F. Van Den Heuvel, "Efficient hough transform for automatic detection of cylinders in point clouds," *Isprs Wg Iii/3, Iii/4*, vol. 3, pp. 60-65, 2005.
- [35] Q.-A. Wang, Y. Dai, Z.-G. Ma, J.-F. Wang, J.-F. Lin, Y.-Q. Ni, W.-X. Ren, J. Jiang, X. Yang, and J.-R. Yan, "Towards high-precision data modeling of shm measurements using an improved sparse bayesian learning scheme with strong generalization ability," *Structural Health Monitoring*, vol. 23, no. 1, pp. 588-604, 2024.
- [36] J.-E. Byun and J. Song, "A general framework of bayesian network for system reliability analysis using junction tree," *Reliability Engineering & System Safety*, vol. 216, p. 107952, 2021.
- [37] Q.-A. Wang, C.-B. Wang, Z.-G. Ma, W. Chen, Y.-Q. Ni, C.-F. Wang, B.-G. Yan, and P.-X. Guan, "Bayesian dynamic linear model framework for structural health monitoring data forecasting and missing data imputation during typhoon events," *Structural Health Monitoring*, vol. 21, no. 6, pp. 2933-2950, 2022.
- [38] S. Rajaraman, D. Caruccio, N. Fung, and C. Hayes, "Fully automatic, unified stereo camera and lidar-camera calibration," in *Automatic Target Recognition XXXI*, vol. 11729, pp. 270-277, SPIE, 2021.
- [39] M. Roy, R. A. Boby, S. Chaudhary, S. Chaudhury, S. D. Roy, and S. K. Saha, "Pose estimation of texture-less cylindrical objects in bin picking using sensor fusion," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2279-2284, IEEE, 2016.
- [40] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (fpfh) for 3d registration," in *2009 IEEE international conference on robotics and automation*, pp. 3212-3217, IEEE, 2009.
- [41] A. Zeng, K.-T. Yu, S. Song, D. Suo, E. Walker, A. Rodriguez,

- and J. Xiao, "Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge," in *2017 IEEE international conference on robotics and automation (ICRA)*, pp. 1386–1383, IEEE, 2017.
- [42] Z. Yaniv, "Random sample consensus (ransac) algorithm, a generic implementation," *Imaging*, 2010.
- [43] A. Jha, M. Soni, and M. Suhaib, "Simulation and kinematic analysis of kuka kr5 arc robot," in *IOP Conference Series: Materials Science and Engineering*, vol. 1149, p. 012005, IOP Publishing, 2021.
- [44] A. Zeng, S. Song, K.-T. Yu, E. Donlon, F. R. Hogan, M. Bauza, D. Ma, O. Taylor, M. Liu, E. Romo, *et al.*, "Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching," *The International Journal of Robotics Research*, vol. 41, no. 7, pp. 690–705, 2022.
- [45] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [46] S. Tang, X. Wang, X. Lv, T. X. Han, J. Keller, Z. He, M. Skubic, and S. Lao, "Histogram of oriented normal vectors for object recognition with a depth sensor," in *Computer Vision—ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5–9, 2012, Revised Selected Papers, Part II 11*, pp. 525–538, Springer, 2013.
- [47] R. Jain, A. Jaju, N. Singh, S. Sharma, and P. K. Pal, "Robotic picking of cylindrical fuel pellets from a boat using 3d range sensor," in *CAD/CAM, Robotics and Factories of the Future: Proceedings of the 28th International Conference on CARs & FoF 2016*, pp. 553–561, Springer, 2016.
- [48] R. Matsumura, Y. Domae, W. Wan, and K. Harada, "Learning based robotic bin-picking for potentially tangled objects," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 7990–7997, IEEE, 2019.
- [49] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3d point cloud based object maps for household environments," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 927–941, 2008.
- [50] Zhuang, C., Li, S., & Ding, H. (2023). Instance segmentation based 6D pose estimation of industrial objects using point clouds for robotic bin-picking. *Robotics and Computer-Integrated Manufacturing*, 82, 102541.
- [51] Zhuang, C., Niu, W., & Wang, H. (2025). Sparse Convolution-Based 6D Pose Estimation for Robotic Bin-Picking With Point Clouds. *Journal of Mechanisms and Robotics*, 17(3), 031007.
- [52] Tao, H., Zheng, Y., Wang, Y., Qiu, J., & Stojanovic, V. (2024). Enhanced feature extraction YOLO industrial small object detection algorithm based on receptive-field attention and multi-scale features. *Measurement Science and Technology*, 35(10), 105023.